USE OF THE LANGUAGE-MEASUREMENT METHOD FOR THE IDENTIFICATION OF PEOPLE IN PHONOSCOPIC EXAMINATIONS

Jerzy DOLECKI, Jacek RZESZOTARSKI

The Faculty of Criminalistics and Special Chemistry, The Agency of Internal Security, Warsaw

ABSTRACT: The Polish school of phonoscopy came into being in the 1960's. One of its first achievements was to work out the methodological foundations for the identification of speakers. Up to then now a practical application of this method had not been presented. The present work is a case study, accompanied by theoretical elements. It encompasses the most significant aspects in the fields of language and phonetics. Despite technological progress and the passage of time, this method is still important in phonoscopic examinations.

KEY WORDS: Phonoscopy; Language-measurement method; Identification of speakers.

Problems of Forensic Sciences, vol. LII, 2002, 108–123
Received 27 November 2002; accepted 30 December 2002

INTRODUCTION

In Polish forensic science, the identification of speakers was first mentioned in publications in the 1960's [11], when the basis for examination methodology was being established at the first phonoscopy laboratory in Poland, attached to the Central Headquarters of the Civic Militia. The method was based on achievements from various fields of science – for example, acoustic phonetics, which became a separate discipline in 1948 [19]. One of the effects of many years of research and the growing numbers of expert's opinions was the development of a language-measurement method for the identification of speakers by a team headed by Stanisław Błasikiewicz. Later publications in forensic science periodicals merely present its theoretical layout [4, 5, 6].

The language-measurement method has been used by phonoscopic experts at the Faculty of Criminalistics and Special Chemistry of the Agency of Inner Security since 1992. The recordings used for expert opinions have an "individual" character: ones of optimum quality and content are rarely subjected to analysis. An analogy can be drawn to the work of fingerprint ex-

perts, who often have to cope with traces which are difficult to reveal due to the characteristic features of the surface.

In their presentation of the practical application of the language-measurement method, the authors will make use of recordings delivered to their laboratory in 2002 as an example.

INTRODUCTION

The essence of language communication is the application of a system of symbols describing the surrounding reality. In the course of the conveyance of spoken communications, the sender conveys not only a chain of words with a literal meaning, but also non-textual information; concerning their emotional state, attitude to the issue in question, level of education and knowledge, outlook on life, etc. This information is also the subject matter of phonoscopic examinations. Linguistic analysis of spontaneous statements is thus conducted in relation to features of speech and not to features of voice.

Speech expresses thoughts and is also the application of linguistic rules and a range of vocabulary. The physicist perceives speech as acoustic "tracks" that are changeable with time in frequency and intensity. This changeability can be illustrated spectrographically [10, 12, 15]. The spectrum of the speech signal reveals a distribution of energy at specific frequencies, i.e. formant structure. Formants – local maximums in the speech sound spectrum – are the effects of the existence of a specific resonance in the larynx [13]. It is due to individual differences in the anatomic structure of the larynx that we are able to differentiate the timbres of voices of various speakers [10].

Formant structures are clearest in the spectrographic images of vowels [3, 10]. Individual vowels enunciated by the same speaker are distinguished by listeners thanks to different formant frequencies, which are the result of articulation, i.e. changes in the speech apparatus [10].

Differences in the location and shape of formants of the same vowels of various speakers are conditioned by:

– anatomical structure of the vocal apparatus;
– phonetic context;
– intonation, and also expression and impression of the statement [3].

As early as the 17th century, a view arose that the genetically conditioned anatomical structure of the supra-larynx resonators determined the scope of the variability of the vocal timbre [19]. In turn, the phonetic surroundings of the allophones of the phoneme "i" in words "świt", "kit", "mit", "nit", and "bit" introduce subtle differences in the spectrum associated with a tiny change in the structure of the larynx and in the efficiency of the cross

striation muscles of the articulation organs [13]. The different intonation, expression and impression of positive, interrogative and imperative sentences also noticeably influences the formant structure [10]. Applying this knowledge to the work of a phonoscopic expert: during collection of the comparative material, the interviewee is directed towards making statements in a specific style and repeats sentences pre-selected for linguistic tests [16]. Material obtained in such a way "illustrates" specific features of the anatomic structure of mouth – pharynx and nasal-pharynx channels and is thus suitable as subject-matter for the measurement of formant frequencies.

The fact that acoustic and phonetic parameters and linguistic features reveal individual differences that result from anatomic, social and cultural conditioning provides the basis for the claim that examination of human speech can be used to categorically distinguish speakers.

PRACTICAL APPLICATION OF THE LANGUAGE MEASUREMENT METHOD

The language measurement-method uses analysis of a statement on two planes: linguistic and phonetic. In order to simplify the description, we are going to present these issues separately.

The requirements which the subject matter of the examination has to meet are not identical. In linguistic analysis, suitable material is voluminous spontaneous speech, whereas in measurement analysis, language tests are used, or, to be more precise – the same allophones of phonemes (the phonetic surroundings are the deciding factor) [14]. The importance of co-articulation is underlined by, for example, the fact that front-tongue consonants heighten the formants of neighbouring vowels "o" and "u", whereas fricatives lower the second formant (F_2) of high vowels [19]. Thus the use of language tests – a sequence of allophones analogous to the evidence material – considerably facilitates the expert's work and shortens the time taken for an opinion to be issued [16].

Language analysis

In every utterance, sets of linguistic features reveal themselves: semantic, syntactic and stylistic ones [8, 9]. The task facing a phonoscopic expert is to extract, from the stream of information, distinctive features that definitely differentiate speakers. In an everyday situation – many telephone users know that some callers start the conversation with "Hello?", and others with "Yes…?", and a few with "Hello, …here". Of course one can not identify a speaker on the basis of such a flimsy premise. In order to make an unequivocal ascertainment, a set of distinctive features must be determined. Analysis of the evidence material in the above mentioned case pointed towards

a series of individual features, connected with a characteristic manner of speaking.

From telephone conversations used as evidence, statements of a man, designated A, have been chosen. Expletives have been deleted from places marked [expl.].

A: Listen! Tomorrow, your deadline is 6 p.m. Get the dough. All right? You've got till 6 tomorrow....
A: ...you can be one hundred percent sure, I tell you, you'll get it back. At six tomorrow. Or else, that's the end of the line...
A: ...and we burn'im, if you don't have the dough...
A: ...You'll know in a moment [expl.].
A: So, how much have you got in the end? How much have you got?! How much?! Come on! [expl.].
A: What [expl.] do you know? – you! What do you know – hey! You wanted it [expl.] – you've got it!
A: You! You know what – a thousand? You've had time to arrange it – ain't you? We're gonna start burning him now [expl.] and... And as you wish....
A: Excuse me, madam – that was the second station, second station that was....
A: ...listen madam...
A: Listen! No car! [expl.]. Give it to me, give me the dough [expl.] and then the car!
A: Are you, [expl.] a kid or something!?

A's utterances have distinct regional features. They also indicate a lack of language fluency, hence the numerous syntactic and stylistic mistakes and very restricted vocabulary supported by many vulgarisms, which, in turn, are heavily emphasised [8, 9]. The above-mentioned examples reveal disorders in syntactic structure of utterances in the form of repetitions (palilalia) [19]. "Colloquial speech" is the best description for the utterances of this speaker. The articulation is rather careless. Despite the fact that the sentences are often formulated contrary to the rules of grammar, the man – in his quieter utterances – is able to convey the message. At moments of high emotional arousal, he has problems with expressing himself, and disorders appear in the pace and rhythm of the speech, tending towards bradylalia (abnormal slowing down) [2, 19]. The man is characterised by a low intellectual level, a further sign of which is the numerous retardation pauses. The expression and impression of the utterances are also characteristic.

The same set of features was revealed in the comparative material. Due to the low intellectual level of the man, recording the comparative sample

was not without problems. The interviewee tried to answer the questions, although sometimes he did not understand them. The questions had to be repeated several times and a commentary had to be provided. The interviewee was often lost for words when answering. He tried not to use expletives, but consequently had a problem replacing them. As a result of this, few spontaneous utterances were obtained. Short, one-word replies prevailed. Even the longest utterances were executed by means of sequences of loose sentences with sporadic use of inter-sentence conjunctions.

Phonetic analysis

After choosing sentences for language tests from the evidence material the expert taking the comparative material tried to obtain comparative utterances with similar intonation, pace, expression and impression. Such material was subjected to the measurement of the formant structure.

The aim of the analysis of the spectral image of speech is to determine the location and the course of the formants of the same allophones. The evidence recording need not be registered in the whole band of audibility. The telephone band (300–3500 Hz) and even narrower, enables to carry out the measurements, whose result might testify to the individuality of a speaker [10].

In the above mentioned case the evidence recordings were executed with the use of a telephone register. With some vowels it was possible to determine the location of even four formants. There was not a problem with the determination of the momentary values of the basic frequency of the larynx tone (F_0). The values of F_0 were concordant in the evidence and the comparative material, which proved that the speakers were in a similar emotional state.

The measurements were carried out with the use of a computerised device DSP Sona-Graph Model made by Kay.

The results of measurements of the position of the corresponding formants in the same allophones were presented in Figures 3–8.

The decisive criterium in the applied method is the dispersion of measurement points in relation to the line of the best fit, which may be characterised by the values of the square of the correlation coefficient (R^2) and the remainder standard deviation (s_y) [1, 7, 18].

The correlation analysis of the formant position of the same allophones in comparative utterance

Table I presents the results of the correlation analysis of the corresponding formants from comparative utterances of speakers, dubbed A, B, C.

In comparisons of formants from several utterances of the same sequence of allophones by the same speakers (the diagonal of the Table I, Figures 3, 5, and 7) the values of the square of correlation coefficient and the remainder standard deviation fulfil the condition {1}:

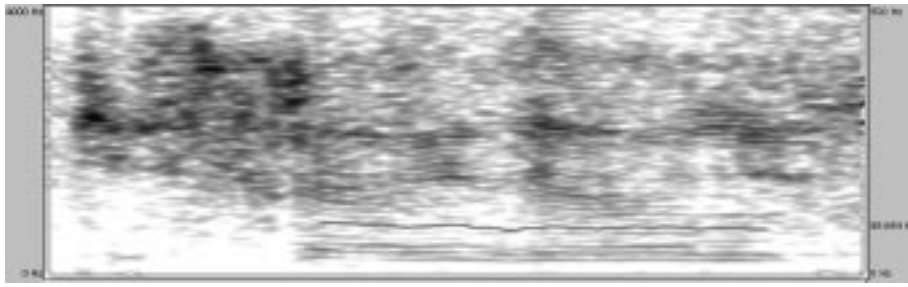

Fig. 1. The basic frequency of larynx tone in the evidence utterance (F_0 mean = 96 Hz).

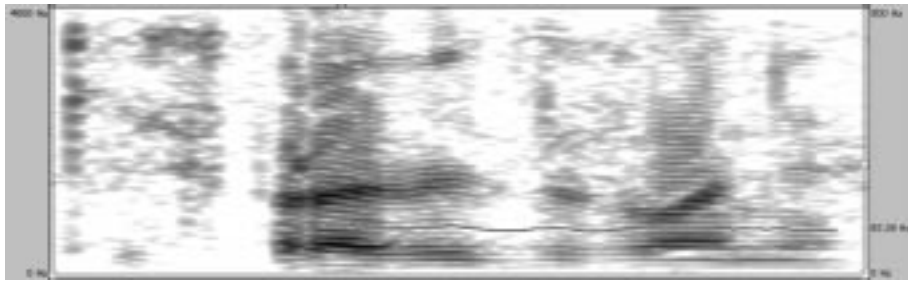

Fig. 2. The basic frequency of larynx tone in the comparative utterance of the man A (F_0 mean = 93 Hz).

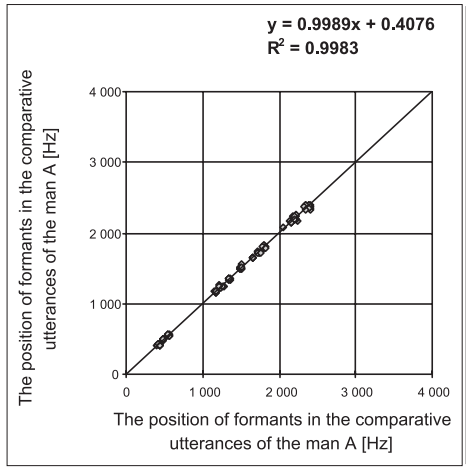

Fig. 3. Correlation of formant positions in comparative utterances of the man A.

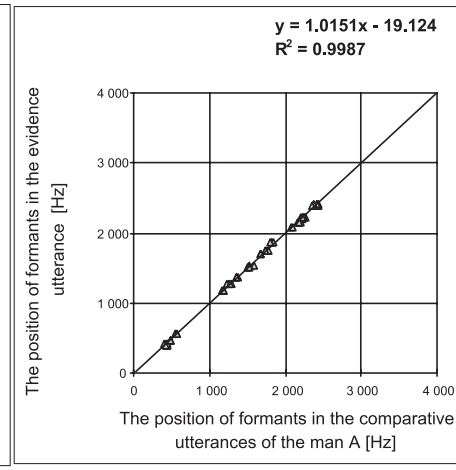

Fig. 4. The correlation of the position of formants in the evidence utterance and the utterance of man A.

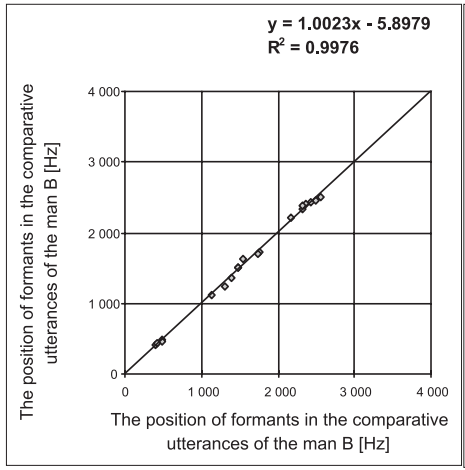

Fig. 5. The correlation of the position of formants in comparative utterances of the man B.

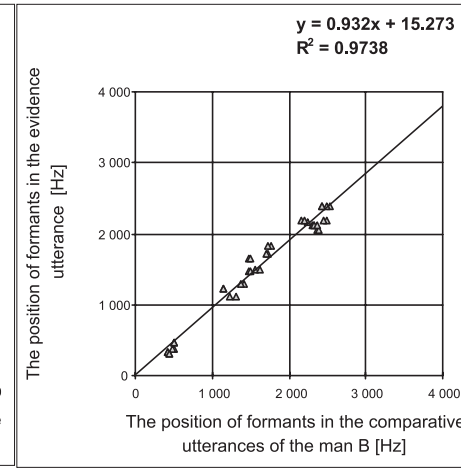

Fig. 6. The correlation of the position of formants in the evidence utterance and the utterance of the man B.

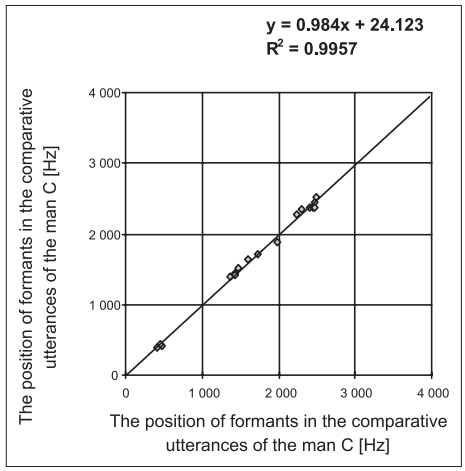

Fig. 7. The correlation of the position of formants in the comparative utterances of the man C.

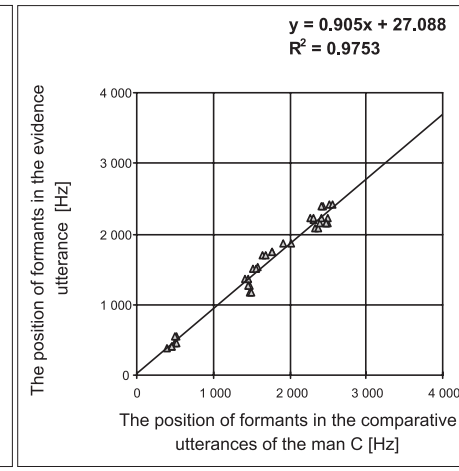

Fig. 8. The correlation of the position of formants in the evidence utterance and the utterance of the man C.

$$R^2 > 0.995 \text{ and } s_y < 50\,. \qquad \{1\}$$

The analogous comparisons of formants for different speakers (cells of Table I not positioned on the diagonal) fulfil the condition {2}:

$$R^2 \leq 0.980 \text{ and } s_y > 100\,. \qquad \{2\}$$

The conditions {1} and {2} define disjoint sets and thus determine, for the considered case, the criteria of the definite identification of speakers on the basis of the formant correlation analysis.

TABLE I. THE RESULTS OF THE CORRELATION ANALYSIS OF A COMPARATIVE UTTERANCE OF MEN A, B, C

Man		A	B	C
A	R^2	0.9983	0.9774	0.976
	s_y	27.5	107	111.9
	Skewness	0.999 ± 0.004	1.055 ± 0.014	1.080 ± 0.015
	Abscissa	4.8 ± 7.1	4.9 ± 22.8	6.2 ± 24.9
B	R^2	0.9774	0.9976	0.98
	s_y	100.2	37	103.4
	Skewness	0.926 ± 0.012	1.002 ±0.012	1.014 ± 0.017
	Abscissa	30.2 ± 21.2	–5.9 ± 21.5	17.1 ± 31.5
C	R^2	0.976	0.98	0.9957
	s_y	102.4	100.9	48.8
	Skewness	0.904 ± 0.012	0.966 ± 0.016	0.984 ± 0.017
	Abscissa	32.0 ± 22.6	17.3 ± 30.7	24.1 ± 37.1

The correlation analysis of the position of formants from the evidence utterance and comparative utterances

Table II and Figures 4, 6 and 8 present the results of the correlation analysis of the corresponding formants from the evidence and comparative utterances of speakers, dubbed A, B and C.

Condition {1} is fulfilled in relation to the evidence utterance and comparative utterances of the man A (Figure 3 and 4).

$$R^2_{(A,evid)} = 0.9987 > 0.995\,, \qquad \{3\}$$

$$s_{y(A,evid)} = 25.7 < 50\,.$$

Condition {2} is fulfilled in relation to the evidence utterance and comparative utterances of speakers B and C (Figures 5, 6, and 7, 8).

The above results support the results of the linguistic analysis that the author of the evidence utterances is the man A.

TABLE II. THE RESULTS OF THE CORRELATION ANALYSIS FOR THE EVIDENCE UTTERANCE (EVID.) AND COMPARATIVE UTTERANCES OF A, B AND C

Man		Evid.
A	R^2	0.9987
	s_y	25.2
	Skewness	1.015 ± 0.005
	Abscissa	–19.1 ± 7.6
B	R^2	0.9738
	s_y	111.3
	Skewness	0.932 ± 0.025
	Abscissa	15.3 ± 45.5
C	R^2	0.9753
	s_y	106.5
	Skewness	0.905 ± 0.024
	Abscissa	21.7 ± 45.8

CONCLUSIONS

1. The language measurement method, taking into consideration a wide range of parameters and linguistic features, provides the basis for the definite identification of speakers.
2. The identification examination of speakers must take into account the fluctuations of features and parameters of speech.
3. The subject matter of linguistic analyses are spontaneous utterances, whereas the subject matter of phonetic analyses are the same sequences of allophones from the evidence and the comparative material.
4. The taking of the comparative material in the form of speech without taking into account the specific features of the evidence material may make it more difficult or even impossible to carry out the identification examination.

SUMMARY

The arguments to apply the language-measurement method should – in the opinion of the authors – convince those, to whom the features and parameters of speech undergo too big fluctuations in order to be accepted as the

safe basis for definite conclusions in an expert's opinion. The contradiction of such a view can be found in the publications of renowned Polish authorities on the subject in the fields which provide theoretical foundations to phonoscopy. [2, 3, 10, 17, 19].

The use of the language-measurement method involves a time-consuming factor of the research resulting in time and cost of the issuing of the opinion.

The basement of the method on the solid scientific foundation and a verified methodology of research allows for the issuing of definite opinions and for a comfortable defence of conclusions in court proceedings.

References:

1 Achnazarowa S., Kafarow W., Optymalizacja eksperymentu, Wydawnictwa Naukowo-Techniczne, Warszawa 1982.
2 Basztura C., Komputerowe systemy diagnostyki akustycznej, Wydawnictwo Naukowe PWN, Warszawa 1996.
3 Basztura C., ródła, sygnały i obrazy akustyczne, Wydawnictwa Komunikacji i Łączności, Warszawa 1988.
4 Błasikiewicz S., Aktualne i przewidywane kierunki rozwoju fonoskopii, Kryminalistyka wczoraj – dziś – jutro, Materiały z Sympozjum z okazji trzydziestolecia Zakładu Kryminalistyki KG MO, Wydawnictwo ZK KG MO, Warszawa 1986.
5 Błasikiewicz S., Osiągnięcia i perspektywy polskiej fonoskopii w walce z przestępczością, *Problemy Kryminalistyki* 1976, nr 123, s. 496–512.
6 Błasikiewicz S., Bednarczyk W., Wpływ transmisji telefonicznej w procesie identyfikacji osób na podstawie analizy mowy, Wydawnictwo ZK KG MO (ZMWF), Warszawa 1972.
7 Czermiński J., Iwasiewicz A., Paszek Z. [i in.], Metody statystyczne, Państwowe Wydawnictwo Naukowe, Warszawa 1974.
8 Grzegorczykowa R., Wprowadzenie do semantyki językoznawczej, Wydawnictwo Naukowe PWN, Warszawa 2001.
9 Grzegorczykowa R., Wykłady z polskiej składni, Wydawnictwo Naukowe PWN, Warszawa 1999.
10 Jassem W., Podstawy fonetyki akustycznej, PWN, Warszawa 1973.
11 Krajowa narada poświęcona pobieraniu materiału porównawczego do badań fonoskopijnych, *Problemy Kryminalistyki* 1969, nr 80–81, s. 634–635.
12 Kuenzel H., Sprechererkennung, Bundeskriminalamt, Heidelberg 1987.
13 Moore B. C. J., Wprowadzenie do psychologii słyszenia, Wydawnictwo Naukowe PWN, Warszawa 1999.
14 Ostaszewska D., Tambor J., Fonetyka i fonologia współczesnego języka polskiego, Wydawnictwo Naukowe PWN, Warszawa 2001.
15 Readings in clinical spectrography of speech, Singular Publishing Group, San Diego, Kay Elemetrics, Pine Brook 1991.

16 Rzeszotarski J., Grzeszczak J., Wybrane aspekty pobrania materiału porównawczego w postaci mowy, [w:] Procesowo-kryminalistyczne czynności dowodowe. Materiały pokonferencyjne, Wyższa Szkoła Policji, Szczytno 2001.
17 Tadeusiewicz R., Sygnał mowy, Wydawnictwa Komunikacji i Łączności, Warszawa 1988.
18 Teoria pomiarów, H. Szydłowski [red.], Państwowe Wydawnictwo Naukowe, Warszawa 1981.
19 Wierzchowska B., Wymowa polska, Państwowe Zakłady Wydawnictw Szkolnych, Warszawa 1971.

ZASTOSOWANIE METODY JĘZYKOWO-POMIAROWEJ DO IDENTYFIKACJI OSÓB W BADANIACH FONOSKOPIJNYCH

Jerzy DOLECKI, Jacek RZESZOTARSKI

WSTĘP

W polskiej kryminalistyce wzmianki o identyfikacji mówców pojawiły się w publikacjach z lat sześćdziesiątych XX wieku [11], kiedy to w pierwszej w kraju pracowni fonoskopii Komendy Głównej MO powstawały podwaliny metodyki badawczej. Oparto się na osiągnięciach wielu dziedzin nauki, choćby wyodrębnionej w 1948 roku fonetyki akustycznej [19]. Jednym z efektów wieloletnich badań i rosnącej liczby wykonywanych ekspertyz było wypracowanie przez zespół pod kierunkiem Stanisława Błasikiewicza językowo-pomiarowej metody identyfikacji mówców. Późniejsze publikacje w periodykach kryminalistycznych przedstawiają jedynie jej zarys teoretyczny [4, 5, 6].

Metoda językowo-pomiarowa jest stosowana przez biegłych fonoskopii Zakładu Kryminalistyki i Chemii Specjalnej Agencji Bezpieczeństwa Wewnętrznego od 1992 roku. W każdej ekspertyzie można mówić o indywidualnej specyfice nagrań. Rzadko poddawane są badaniom nagrania, które cechuje optymalna jakość i zawartość merytoryczna. Występuje tu analogia chociażby do pracy ekspertów daktyloskopii, którzy często zmagają się z trudnymi do ujawnienia – ze względu na specyfikę podłoża – śladami.

Prezentując praktyczne zastosowanie metody językowo-pomiarowej, autorzy posłużą się przykładem nagrań, które trafiły do ich pracowni w 2002 roku.

WPROWADZENIE

Istotą porozumiewania językowego jest stosowanie systemu znaków opisujących otaczającą rzeczywistość. W trakcie przekazywania komunikatów słownych nadawca przekazuje nie tylko właściwą treść przypisaną do ciągu wyrażeń, ale też szereg informacji pozatekstowych: o swoim stanie emocjonalnym, stosunku do omawianego zagadnienia, wykształceniu, zasobie wiedzy, światopoglądzie itp. Również te informacje są przedmiotem badań fonoskopijnych. Analiza lingwistyczna wypowiedzi spontanicznych jest więc przeprowadzana w odniesieniu do cech mowy, nie głosu.

Mowa wyraża myśli, jest również realizacją zasad językowych i zasobu słownictwa. Fizyk postrzega mowę jako przebiegi akustyczne zmienne w czasie w dziedzinie częstotliwości i natężenia. Zmienność ta może być ilustrowana spektrograficznie [10, 12, 15]. Widmo sygnału mowy ukazuje rozkład energii w poszczególnych częstotliwościach, czyli strukturę formantową. Formanty – maksima lokalne w widmie dźwięku mowy – są efektem istnienia określonego rezonansu w kanale głosowym [13]. Różnicom osobniczym budowy anatomicznej kanału głosowego zawdzięczamy rozróżnianie barwy głosu różnych mówców [10].

Struktury formantowe są najbardziej przejrzyste w obrazach spektrograficznych samogłosek [3, 10]. Poszczególne samogłoski wymawiane przez tego samego mówcę są odróżniane przez słuchaczy dzięki innym częstotliwościom formantowym, będącym wynikiem artykulacji czyli zmian układu aparatu mowy [10].

Odmienności w położeniu i kształcie formantów tych samych samogłosek różnych mówców są warunkowane przez:

– budowę anatomiczną aparatu mowy;
– kontekst fonetyczny;
– intonację oraz ekspresję i impresję wypowiedzi [3].

Już w XVII wieku pojawił się pogląd, że uwarunkowana genetycznie budowa anatomiczna rezonatorów nadkrtaniowych określa ramy zmienności barwy głosu [19]. Z kolei otoczenie fonetyczne alofonów fonemu „i" w słowie „świt", „kit", „mit", „nit" i „bit" wprowadza subtelne różnice w widmie związane z nieznaczną zmianą układu kanału głosowego i sprawnością mięśni poprzecznie prążkowanych narządów artykulacji mowy [13]. Odmienna intonacja, ekspresja i impresja zdania oznajmującego, pytającego czy rozkazującego także zauważalnie wpływa na strukturę formantową [10]. Przekładając to na specyfikę pracy eksperta fonoskopii: w trakcie pobierania materiału porównawczego indagowany jest naprowadzany na konkretny styl wypowiedzi i powtarza wybrane – do testów językowych – zdania [16]. Tak otrzymany materiał jest wykładnią specyfiki anatomicznej kanału ustno-gardłowego i nosowo-gardłowego, jest więc właściwy do przeprowadzenia pomiarów częstotliwości formantowych.

Fakt, że parametry fonetyczno-akustyczne i cechy lingwistyczne uwidaczniają indywidualność osobniczą wynikającą z uwarunkowań anatomicznych, społecznych i kulturowych, leży u podstaw twierdzenia, że badanie mowy ludzkiej pozwala kategorycznie odróżniać mówców.

PRAKTYCZNE ZASTOSOWANIE METODY JĘZYKOWO-POMIAROWEJ

Metoda językowo-pomiarowa zakłada analizę wypowiedzi na dwóch głównych płaszczyznach: językowej i fonetycznej. Dla uproszczenia opisu przedstawiamy te zagadnienia oddzielnie.

Wymogi, które musi spełniać przedmiot badań, nie są jednakowe. W analizie językowej odpowiednim materiałem są obszerne wypowiedzi spontaniczne, natomiast w analizie pomiarowej testy językowe, a ściślej – te same alofony fonemów (decyduje otoczenie fonetyczne) [14]. Znaczenie koartykulacji podkreśla chociażby przykład podwyższania przez spółgłoski przedniojęzykowe formantów sąsiadujących z nimi samogłosek „o", „u", natomiast spółgłoski drżące obniżają drugi formant (F_2) samogłosek wysokich [19]. Tak więc dysponowanie testami językowymi – ciągiem alofonów analogicznym do materiału dowodowego – znacznie ułatwia pracę biegłym i skraca czas wydania opinii [16].

Analiza językowa

W każdej wypowiedzi ujawniają się zespoły cech językowych: semantyczne, składniowe i stylistyczne [8, 9]. Zadaniem eksperta fonoskopii jest wyekstrahowanie ze strumienia informacji cech dystynktywnych, kategorycznie odróżniających po-

szczególnych mówców. W przełożeniu na sytuację z życia codziennego: niejeden użytkownik telefonu wie, że część abonentów rozpoczyna rozmowę od „halo?", inni od „tak, słucham?", nieliczni natomiast od „słucham uprzejmie?" Oczywiście nie można na tak skromnej przesłance opierać sądów o niepowtarzalności. Do jednoznacznych stwierdzeń uprawnia określenie zespołu cech dystynktywnych. Analiza materiału dowodowego we wspomnianej już sprawie pozwoliła wskazać szereg cech indywidualnych, związanych z charakterystycznym sposobem mówienia.

Z dowodowych rozmów telefonicznych wybrano wypowiedzi mężczyzny oznaczonego literą A. Wulgaryzmy usunięto z miejsc oznaczonych [wulg.]

A: Słuchaj! Jutro, do godziny osiemnastej masz ostatni dzień. Zbieraj kasę. W porządku? Do osiemnastej jutro...
A: ...sto procent, mówię ci, odzyskasz. O osiemnastej jutro. Bo koniec karty...
A: ...i palim go, jak nie będziesz miał kasy...
A: ...bo zaraz się dowiesz [wulg.] gdzie stoi, gdzie będzie hajcował się będzie...
A: A ile masz w końcu, [wulg.] no ile masz?! [wulg.] No, ile!? Mów! [wulg.]
A: Co ty [wulg.] możesz mówić – ty! Co ty możesz mówić – no! Chciałeś [wulg.] – to masz!
A: Ty! Wiesz co – tysiąc? Miałeś czas do załatwienia – nie? Zaraz będziem palić [wulg.] go i o! Jak sobie chcesz...
A: ...proszę panią – stacyjka była druga była stacyjka...
A: ...słucha pani...
A: Słuchaj! Nie ma samochód! [wulg.] Dajesz, dajesz kasę, [wulg.] potem samochód! [wulg.] Co ty, [wulg.] dzieciak mały jesteś?!

Wypowiedzi mężczyzny A noszą wyraźne znamiona regionalne. Świadczą również o braku biegłości językowej, czego wyrazem są liczne błędy syntaktyczno-stylistyczne i bardzo ograniczone słownictwo podpierane licznymi wulgaryzmami; te z kolei są silnie akcentowane [8, 9]. W przytoczonych przykładach dostrzec można zaburzenia struktury syntaktycznej wypowiedzi w postaci powtórzeń wyrazów (palilalia) [19]. Dla wytworów fonacji tego mówcy właściwym jest określenie „mowa potoczna". Artykulacja jest dość niedbała. Mimo, że zdania są niekiedy formułowane niezgodnie z regułami gramatyki, mężczyzna – w spokojnych wypowiedziach – radzi sobie z przekazywaniem konkretnych treści. W chwilach silnego wzburzenia emocjonalnego ma problemy z wysłowieniem się, a także pojawiają się u niego zaburzenia tempa i rytmu mowy w kierunku bradylalii (nienormalne zwalnianie) [2, 19]. Mężczyznę cechuje niski poziom intelektualny, czego kolejnym przejawem są liczne pauzy retardacyjne. Charakterystyczna jest ekspresja i impresja wypowiedzi.

Taki sam zestaw cech uwidocznił się także w materiale porównawczym. Z powodu niskiego poziomu intelektualnego podejrzanego czynność pobrania wypowiedzi porównawczych przeprowadzono nie bez problemów. Indagowany starał się odpowiadać na pytania, jednak nie zawsze rozumiał ich treść. Pytania musiały być zadawane kilkakrotnie i opatrywane komentarzem. Pytanemu często brakowało słów przy formułowaniu odpowiedzi. Starał się nie używać wulgaryzmów, więc miał widoczny problem z ich zastąpieniem. W efekcie uzyskano niewiele wypowiedzi spontanicznych, przeważały odpowiedzi krótkie, jednowyrazowe. Nawet najdłuższe

wypowiedzenia były realizowane ciągami luźnymi zdań ze sporadycznymi nawiązaniami międzyzdaniowymi.

Analiza fonetyczna

Po wybraniu z materiału dowodowego zdań do testów językowych, biegły przeprowadzający czynność pobrania materiału porównawczego, dążył do uzyskania wypowiedzi porównawczych o zbliżonej intonacji, tempie, ekspresji i impresji. Tak uzyskany materiał posłużył do pomiarów struktury formantowej.

Celem przeprowadzenia analizy obrazu widmowego mowy jest ustalenie położenia i przebiegu formantów tych samych alofonów. Nagranie dowodowe nie musi być zarejestrowane w całym paśmie słyszalności. Pasmo telefoniczne (300–3500 Hz), a czasami nawet węższe, umożliwia przeprowadzenie pomiarów, których wynik może świadczyć o indywidualności mówcy [10].

W omawianym przypadku nagrania dowodowe zostały utrwalone z wykorzystaniem rejestratora telefonicznego. Dla niektórych samogłosek możliwe było wyznaczenie położenia nawet czterech formantów. Nie istniały problemy z wyznaczeniem wartości chwilowych częstotliwości podstawowej tonu krtaniowego (F_0). Wartości F_0 wykazały zgodność w materiale dowodowym i porównawczym, co potwierdziło, że rozmówcy znajdowali się w podobnym stanie emocjonalnym.

Pomiary zostały przeprowadzone z wykorzystaniem skomputeryzowanego urządzenia DSP Sona-Graph Model 5500 firmy Kay.

Wyniki pomiarów położenia odpowiadających sobie formantów w tych samych alofonach przedstawiono w układzie współrzędnych prostokątnych na rycinach 3–8. Kryterium rozstrzygającym w stosowanej metodzie jest rozproszenie punktów pomiarowych względem linii najlepszego dopasowania, które można scharakteryzować wartościami kwadratu współczynnika korelacji (R^2) i resztkowego odchylenia standardowego (s_y) [1, 7, 18].

Analiza korelacyjna położenia formantów tych samych alofonów w obrębie wypowiedzi porównawczych

W tabeli I przedstawiono wyniki analizy korelacyjnej odpowiadających sobie formantów z wypowiedzi porównawczych mówców oznaczonych literami A, B i C.

W zestawieniach formantów z kilku wypowiedzeń tych samych ciągów alofonów przez tych samych mówców (przekątna tabeli I, ryciny 3, 5 i 7) wartości kwadratu współczynnika korelacji i resztkowego odchylenia standardowego spełniają warunek {1}:

$$R^2 > 0.995 \text{ and } s_y < 50 . \qquad \{1\}$$

Analogiczne zestawienia formantów dla różnych mówców (komórki tabeli I nie leżące na przekątnej) spełniają warunek {2}:

$$R^2 \leq 0.980 \text{ and } s_y > 100 . \qquad \{2\}$$

Warunki {1} i {2} definiują zbiory rozłączne, a tym samym określają, dla rozważanego przypadku, kryteria kategorycznej identyfikacji mówców na podstawie korelacyjnej analizy formatowej.

Analiza korelacyjna położenia formantów z wypowiedzi dowodowej i wypowiedzi porównawczych

W tabeli II oraz na rycinach 4, 6 i 8 przedstawiono wyniki analizy korelacyjnej odpowiadających sobie formantów z wypowiedzi dowodowej i wypowiedzi porównawczych mówców, oznaczonych literami A, B i C.

Warunek {1} jest spełniony w zestawieniu wypowiedzi dowodowej z wypowiedziami porównawczymi mężczyzny A (ryciny 3 i 4):

$$R^2_{(A,Dow)} = 0.9987 > 0.995\,, \qquad \{3\}$$

$$s_{y\,(A,Dow)} = 25.7 < 50\,.$$

Zestawienia wypowiedzi dowodowej i wypowiedzi porównawczych mówców B i C spełniają warunek {2} (ryciny 5 i 6 oraz 7 i 8).

Powyższe rezultaty potwierdzają wynik analizy językowej, że autorem wypowiedzi dowodowej jest mężczyzna oznaczony literą A.

WNIOSKI

1. Metoda językowo-pomiarowa, uwzględniając szerokie spektrum cech i parametrów występujących w mowie, daje podstawy do kategorycznej identyfikacji mówców.
2. Badania identyfikacyjne mówców muszą uwzględniać fluktuacje cech i parametrów mowy.
3. Przedmiotem analizy lingwistycznej są spontaniczne wypowiedzi, natomiast analizy fonetycznej te same ciągi alofonów z materiału dowodowego i porównawczego.
4. Pobranie materiału porównawczego w postaci mowy bez uwzględnienia specyfiki materiału dowodowego może utrudnić lub wręcz uniemożliwić przeprowadzenie badań identyfikacyjnych.

PODSUMOWANIE

Argumenty przemawiające za praktycznym zastosowaniem metody językowo-pomiarowej powinny – w odczuciu autorów – przekonać tych, dla których cechy i parametry mowy podlegają zbyt dużym fluktuacjom, aby można było opierać na nich kategoryczne wnioski opinii kryminalistycznej. Zaprzeczenie takiego poglądu znaleźć można choćby w publikacjach wybitnych polskich autorytetów naukowych w dziedzinach, które fonoskopii dostarczają podstaw teoretycznych [2, 3, 10, 17, 19].

Stosowanie metody językowo-pomiarowej implikuje znaczną pracochłonność badań, przekładającą się na czas i koszt wydania opinii.

Oparcie metody identyfikacji mówców na solidnych podstawach naukowych i zweryfikowanej metodyce badań pozwala wydawać kategoryczne opinie i skutecznie bronić zawarte w nich wnioski w postępowaniu sądowym.